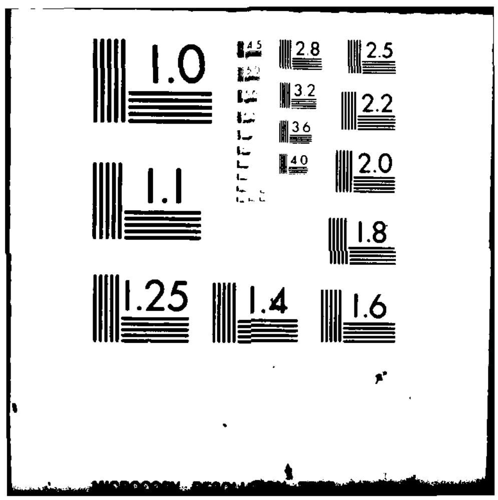
AD-A107 443   NAVAL POSTGRADUATE SCHOOL MONTEREY CA   F/G 9/4
ESTIMATION OF SENSOR DETECTION PROBABILITIES WITH DATA FROM CON--ETC(U)
AUG 81 D R BARR

UNCLASSIFIED NPS55-81-014   NL

1 of 1
AD
A107443

END
DATE
FILMED
12-81
DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

NPS55-81-014

NAVAL POSTGRADUATE SCHOOL

Monterey, California

ESTIMATION OF SENSOR DETECTION PROBABILITIES WITH DATA FROM CONCURRENT SENSORS

by

D. R. Barr

August 1981

Approved for public release; distribution unlimited

Prepared for:

Chief of Naval Research (Code 521)
800 North Quincy Street
Arlington, Virginia 22217

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral J. J. Ekelund
Superintendent

D. A. Schrady
Acting Provost

This report was prepared by:

Donald R. Barr

Donald R. Barr, Professor
Department of Operations Research

Reviewed by

K. T. Marshall

K. T. Marshall, Chairman
Department of Operations Research

William M. Tolles

William M. Tolles
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55-81-014 — 2. GOVT ACCESSION NO. ADA107443	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) ESTIMATION OF SENSOR DETECTION PROBABILITIES WITH DATA FROM CONCURRENT SENSORS.	5. TYPE OF REPORT & PERIOD COVERED Technical rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) D. R./Barr	8. CONTRACT OR GRANT NUMBER(s) N0001481WRP1067
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940	12. REPORT DATE August 1981
	13. NUMBER OF PAGES ~~25~~ 27
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Chief of Naval Research (Code 521) 800 North Quincy Street Arlington, VA 22217	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.	
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)	
18. SUPPLEMENTARY NOTES	
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) detection probability; sensor data; intraclass correlation; method of moments; least squares; maximum likelihood	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

When several sensors are concurrently scanning the same domain for signals, varying numbers of sensors may detect each signal. On some occasions, a signal may not be detected by any of the receivers. Using detection data collected from all the receivers over a period of scanning, it is possible to estimate the total number of signals that occurred in that period (including those that were not detected at all), as well as the detection probabilities for the individual receivers. Several estimators for

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

251450

these quantities are developed, in the contexts of several models concerning the signal generation process and the receiver behavior.

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

ESTIMATION OF SENSOR DETECTION PROBABILITIES WITH DATA FROM CONCURRENT SENSORS

by

D. R. Barr

Abstract

When several sensors are concurrently scanning the same domain for signals, varying numbers of sensors may detect each signal. On some occasions, a signal may not be detected by any of the receivers. Using detection data collected from all the receivers over a period of scanning, it is possible to estimate the total number of signals that occurred in that period (including those that were not detected at all), as well as the detection probabilities for the individual receivers. Several estimators for these quantities are developed, in the contexts of several models concerning the signal generation process and the receiver behavior.

Key words: detection probability, sensor data, intraclass correlation, method of moments, least squares, maximum likelihood.

1. Introduction

The problem we consider concerns estimation of performance characteristics of sensors, based on historical sensor reception data. We shall imagine the sensors are "receivers", passively listening for "signals". If only one receiver is turned on, an operator might believe his receiver is detecting all of the signals that occur over time; he would have no data indicating otherwise. If a second receiver were turned on, to operate concurrently with the first, the operator might notice that sometimes one of the receivers detects a signal the other receiver does not, and vice versa. This implies there are signals occurring that neither receiver detects, and the detection probabilities for the receivers are less than unity.

Suppose k receivers operate concurrently and a record is kept of the signals detected by each one. It is assumed that the times of detection, and perhaps the nature of the signals detected, makes it possible to determine when detections on several receivers are detections of a common signal. It is convenient to imagine the data collected from the detection history with k receivers are in the form of k-dimensional vectors, one vector for each signal detected. The components of such a data vector are zeros (for "non-detect") and ones (for "detect"), each component corresponding to a particular receiver. Thus, for example, the sum of components of a data vector for a given signal would be the number of receivers that detected that signal. Note this sum must be at least 1, since

otherwise one would have no knowledge a signal occurred and hence there would be no related data vector. Of course, signals may occur that no receiver detects; we will be interested in estimating how frequently this occurs.

We wish to use the set of data vectors observed over some period of time to estimate detection probabilities of the receivers, and perhaps to estimate characteristics of the signal process, such as the distribution of signal strength and the total number of signals occurring in the period in question. Various estimation principles, such as maximum likelihood, method of moments and least squares, can be used in the context of various models for the signal detection process. Thus there are a number of candidate models and estimators which could be used for the problem we have described. The purpose of this report is to develop several estimators that could be used for this problem; a later report will compare their behavior under varying conditions of the signal process. We also give a brief review of the literature related to this problem and discuss associated topics such as intraclass correlation of the receivers and possible effects of dependence among the receivers.

2. Related Literature

The problem we consider has some similarities to the capture-tag-recapture methods used to estimate wildlife populations. The initial capture of an animal is analogous to the detection of a signal by one receiver; subsequent recapture of that animal is analogous to detection of the same signal by a

second receiver. The presence of a tag on a recaptured animal makes it possible to determine whether the two "receivers" have "detected" a common "signal". There is a well established literature on this and related methods of wildlife population estimation, and these methods might be useful for the case of $k = 2$ receivers. (See, for example, Cormack (1972) for a review of capture-recapture methods and a good bibliography on the subject.)

Applications of the methods we discuss can be made in a wide range of situations. It could be used to estimate the number of errors in computer software which is checked by several independent testors ("receivers"). The models have been used in connection with visual scanning of film containing images of particles occurring in a bubble chamber, with several film scanners acting as "receivers" (Sanathanan (1972)). The application to receivers detecting signals was investigated by Knorr (1979); one of the methods we discuss below is developed in Knorr's report. I am indebted to Professor Knorr for suggesting this problem and for providing suggestions and support during the course of my work on the problem.

3. The Model

Suppose signal strength varies randomly from signal to signal, so that receiver i detects signal j with probability π_{ij}, independent of other signals. Let the random variable V_j denote an index of the "strength" or "detectability" of the jth signal, where $0 \leq V_j \leq 1$ with

probability 1. Assume the probability π_{ij} is related to signal strength v_j through some receiver specific function τ_i. That is, $\pi_{ij} = \tau_i(v_j)$*, where v_j is the outcome on V_j. We suppose the receivers are <u>conditionally</u> independent, in the sense that, given $V = v$, the joint probability of a given vector of detect-nondetect components is the product of the marginal probabilities for each individual receiver. Symbolically,

$$(1) \qquad P[Z = z \mid V = v] = \prod_{i=1}^{k} P[Z^i = z^i \mid V = v]$$

$$= \prod_{i=1}^{k} \tau_i(v)^{z^i} (1 - \tau_i(v))^{1-z^i} ,$$

where Z is a random k-vector whose components Z^i represent the individual detect/non-detect events that will occur on the $i\underline{th}$ receiver and z and z^i denote points in the sample spaces of Z and Z^i, respectively. Thus Z^i is a Bernoulli random variable and Z is a vector of conditionally (given V) independent Bernoulli components. For an arbitrary (future) signal strength V, let Π_i denote $\tau_i(V)$ and $\pi_i = \tau_i(v)$ (i.e., drop the second subscript on π_{ij} when it carries no information).

Let S denote the sample space of Z, so S contains 2^k k-dimensional vectors of zeros and ones. (We include the zero vector, $\underline{0}$, in S, even though it cannot be observed as

---

*A list of the symbols used in this report is given in the Appendix.

an outcome on Z, for reasons described above.) The unconditional probability of observing an outcome $z \varepsilon S$ is obtained by integrating (1) with respect to the marginal distribution F of V:

$$(2) \qquad p^z = P[Z = z] = \int_0^1 P[Z = z|v] \, dF(v)$$

$$= E\left[\prod_{i=1}^{k} \tau_i(V)^{z^i}(1 - \tau_i(V))^{1-z^i}\right],$$

where "E" denotes expectation with respect to V. By the conditional independence assumption (1), the unconditional marginal probability the $i\underline{th}$ receiver detects a randomly selected signal is

$$(3) \qquad \pi^i = P[Z^i = 1] = \int_0^1 \tau_i(v) \, dF(v) \; .$$

One could consider several forms of the detection functions τ_i for receiver i; i=1,2,...,k, including the following:

(a) (constant signal strength) $\tau_i(v) \equiv \ell_i$ for all i. This is the case studied by Knorr.

(b) (equal receivers) $\tau_i(v) \equiv v$ for all i.

(c) (receiver threshold) $\tau_i(v) = \begin{cases} 1 & \text{if } v \geq \ell_i \\ 0 & \text{if } v < \ell_i \end{cases}$,

where ℓ_i is the threshold limit for receiver i .

(d) (fuzzy threshold) $\tau_i(v) = \ell_i \cdot v$, where ℓ_i denotes the proportion of signals of strength v that receiver i would detect.

By taking special forms of the distribution of V and the constants ℓ_i , it can be seen that (a), (b) and (c) are all special cases of (d). Thus, (a) results with V degenerate at 1 in (d); (b) results with $\ell_i \equiv 1$ in (d); (c) results with $\ell_i \equiv 1$ and V distributed Bernoulli in (d). Thus, we shall consider only the case (d) with various conditions on the ℓ_i's and distributions for V .

For each $z \in S$, let n^z denote the number of times z occurs in the data set (i.e., the number of times this particular set of detects and non-detects, among the k receivers, occurred during the time period in question). Let n be the 2^k-dimensional vector whose components are the n^z's in some specified order*, and let N be the random vector with outcome n , where n varies over N's sample space, say T. It follows that N is distributed multinomial with "number of trials" parameter s and "type z outcome" probability vector p whose components p^z are given in eq. (2). Note s is the <u>total</u> number of signals occurring, including the undetected signals, so s is unobservable. Symbolically, we summarize this by writing

---

*A convenient method of indexing the z's is to regard the components of z as forming a binary number, then indexing z with the decimal equivalent of that binary number. Thus, for example, $0 = (0,0,0,\ldots,0)$ is indexed by 0 , $(0,0,\ldots,0,1,0,1)$ is indexed by 5 (since $101_{binary} = 5_{decimal}$), and so on.

$$N \sim M_{2^k}(s,p) \;,$$

where the 2^k components of p are given by

$$(2') \qquad p^z = \int_0^1 \prod_{i=1}^{k} \left[(\ell_i \cdot v)^{z^i} (1 - \ell_i \cdot v)^{1-z^i} \right] dF(v) \;,$$

and may depend on the ℓ_i's as well as parameters in $F(v)$. We use this multinomial formulation in the maximum likelihood estimation approaches discussed below.

4. Intraclass Correlation

It may be the case that the k receivers operate physically independently of one another and yet, because they all are subjected to the same signal "stimulus" at a given instant of signal occurrence, the receiver detection outputs are statistically dependent. This is analogous to intraclass correlation in experiments having random effects (see Scheffé (1959)). We are assuming receivers i and j detect a given common signal having strength v with probability $(\ell_i \cdot v)(\ell_j \cdot v)$. Recall z^i and z^j are Bernoulli random variables whose outcomes (0's or 1's) designate the respective receiver i and j detection or nondetection of the given signal in question. We shall show the correlation between z^i and z^j can be (and generally is) positive, so the receiver detections are not in general independent.

Now

$$(4) \qquad \mathrm{Cov}(Z^i, Z^j) = E\,(Z^i Z^j) - E(Z^i)E(Z^j)$$

$$= \ell_i \ell_j \int_0^1 v^2 dF(v) - \ell_i \ell_j \left(\int_0^1 v dF(v) \right)^2$$

$$= \ell_i \ell_j \sigma_V^2 ,$$

where σ_V^2 is the variance of signal strength, V. Similarly, the variance of Z^i is

$$V(Z^i) = E_V(V(Z^i|V)) + V_V(E(Z^i|V))$$

$$= E(\ell_i V(1-\ell_i V)) + V(\ell_i V)$$

$$= \ell_i E(V) - \ell_i^2 E(V^2) + \ell_i^2 V(V)$$

$$= \ell_i \mu_V (1-\ell_i \mu_V) ,$$

where μ_V is the mean signal strength, with a similar expression for the variance of Z^j. It follows that the correlation of Z^i and Z^j is

$$(5) \qquad \rho(Z^i, Z^j) = \frac{\sigma_V^2 \sqrt{\ell_i \ell_j}}{\mu_V \sqrt{(1-\ell_i \mu_V)(1-\ell_j \mu_V)}} = \geq 0 .$$

If signal strength is constant, say degenerate at 1, then $\sigma_V^2 = 0$ and this intraclass correlation is zero. On the other

hand, if signal strength varies among signals, then $\rho > 0$ and the receiver outputs cannot be independent. It is useful to note from (3) and (4) the probability that receivers i and j both detect a given signal is

$$(6) \qquad E(Z^i \cdot Z^j) = \ell_i \ell_j E(V^2) = \ell_i \ell_j (\sigma_V^2 + \mu_V^2)$$
$$= \ell_i \ell_j \sigma_V^2 + \pi^i \pi^j = \pi^i \pi^j \left(1 + \frac{\sigma_V^2}{\mu_V^2}\right),$$

the latter quantity being related to the signal to noise ratio of the signal, V.

As a specific numeric example, suppose V has a beta distribution with parameters α and 1, so that $F(v) = v^\alpha$ for $0 \le v \le 1$ and $\alpha > 0$. Then $\sigma_V^2 = \alpha/(\alpha+1)^2(\alpha+2)$ and $\mu_V = \alpha/(\alpha+1)$, so using (5), we have for this case

$$\rho(Z^i, Z^j) = \sqrt{\ell_i \ell_j}/(\alpha+2)\sqrt{(1 + \alpha(1-\ell_i))(1 + \alpha(1-\ell_j))} .$$

If $\alpha = 1$, so V is distributed uniform over $(0,1)$ then $0 \le \rho \le 1/3$, depending on the values of ℓ_i and ℓ_j between 0 and 1. ($\rho \simeq .33$ if the ℓ's are 1; $\rho \simeq .1$ if the ℓ's are .5). If $\alpha = 100$ (say), the distribution of V is nearly degenerate at 1 ($\mu_V = .99$ and $\sigma_V^2 = 9.6 \times 10^{-5}$), and the intraclass correlation is nearly zero ($0 \le \rho \le .01$).

In a model that assumes (and exploits) independence of the receivers, the presence of positive correlation can be

expected to degrade the estimates of π and s (the vector of receiver detection probabilities and number of signals present, respectively). The nature of the degradation depends on the particular situation and may be difficult to determine. It can be expected that the variance of the estimators, and the mean square error of the estimators, will increase if a model is used which ignores the receiver dependence if it is in fact present. It should be possible to get some idea of the strength of correlation among receivers by examining certain residuals. For example, suppose π and s are estimated by $\hat{\pi}$ and $\hat{s}$, using estimators which are derived assuming receiver independence, and suppose $\hat{\pi}$ and $\hat{s}$ are used to estimate the expected number of signals that should have been detected by exactly j receivers, for $j = 0,1,2,\ldots,k$. Define the jth residual by

(7) $R(j)$ = (observed number of signals detected by j receivers) − (expected number of signals detected by j receivers).

If one finds $R(j)$ tends to be positive for j's near 0 and near k, while $R(j)$ tends to be negative for intermediate j's, this suggests $\rho > 0$. In such a case, the quality of $\hat{\pi}$ and $\hat{s}$ may be questionable, and it might be prudent to use alternate estimators which accommodate dependence among the receivers.

5. Estimation Approaches

We shall discuss several approaches to deriving estimators for π and s. Most of these require numerical minimization techniques, which can be carried out with any one of several commonly available programs, such as GRG (Lasdon, et al (1975)). The approaches fall into three categories: Method of Moments, Least Squares and Maximum Likelihood. In all cases, $\sigma_V^2 = 0$ is assumed in the estimation model, except as noted in the Maximum Likelihood approach.

(a) Method of Moments Estimators. Let T^i denote the number of signals detected by receiver i. Then T^i is binomially distributed with parameters s and π^i, so $E(T^i) = s \cdot \pi^i$; $i = 1,2,\ldots,k$. Thus k equations in the $k + 1$ unknowns $s, \pi^1, \pi^2,\ldots,\pi^k$ are obtained by setting $t^i = \hat{s}\hat{\pi}^i$; $i = 1,2,\ldots,k$, where t^i is the observed outcome on the random variable T^i. To obtain a $k + 1$ st equation, let

$$T^* = \sum_{i<j} \text{(number of signals detected simultaneously by receivers } i \text{ and } j)$$

so

$$E(T^*) = s \sum_{i<j} \pi^i\pi^j \left(1 + \frac{\sigma_V^2}{\mu_V^2}\right) = s \sum_{i<j} \pi^i\pi^j ,$$

assuming $\sigma_V^2 = 0$. Set $t^* = \hat{s} \sum_{i<j} \hat{\pi}^i\hat{\pi}^j$ and, together with the k equations $t^i = \hat{s}\hat{\pi}^i$; $i = 1,2,\ldots,k$, solve for $\hat{s}$ and $\hat{\pi}^i$ to obtain

8) $$\hat{s} = \left[\sum_{i<j} t^i t^j\right] / t^* \; ; \; \hat{\pi}^i = t^i t^* / \sum_{i<j} t^i t^j$$

(b) Least Squares Estimators. First, let us consider the method developed by Knorr (1979). Let A denote the event $[Z \neq \underline{0}]$ that a given signal is detected by at least one receiver. Recall t^i denotes $\sum z^i$ (sum over all data vectors z), the total number of signals detected by receiver i ; let $t^{\cdot}$ denote the number of such data vectors (i.e., the total number of signals detected). Now π^i might be estimated by

(9) $$\hat{\pi}^i = \frac{t^i}{t^{\cdot}} \cdot P(A)$$

if $P(A)$ were known, since $t^i/t^{\cdot}$ estimates the relative frequency of the event [receiver i detects] among those occasions for which the event A = [some receiver detects] occurred. Similarly, let $Z^{\cdot} = \sum_{i=1}^{k} Z^i$ denote the number of detections made of a given signal, let I^j denote the number of data vectors for which $z^{\cdot} = j$, and let $p_{Z^{\cdot}}$ be the mass function of $Z^{\cdot}$ (positive at $z^{\cdot} = 0,1,\ldots,k$). Now $p_{Z^{\cdot}}(j)$, the probability of exactly j detections of a given signal, might be estimated (for $j > 0$) by:

(10) $$\hat{p}_{Z^{\cdot}}(j) = \frac{I^j}{t^{\cdot}} \cdot P(A)$$

if $P(A)$ were known, since again, $I^j/t^\bullet$ is a relative frequency which estimates the conditional probability $P[Z^\bullet = j|A]$. For the case $j = 0$,

$$\hat{p}_{Z^\bullet}(0) = 1 - P(A) . \tag{11}$$

Now consider $P(A)$ to be a parameter which can vary over the interval $(0,1)$. For each value of $P(A)$, and a given data set determining t^i, $t^\bullet$ and I^j, one can calculate $\hat{p}_{Z^\bullet}(j)$ using (10) and (11). Alternatively, one can estimate the corresponding probabilities by first estimating the $\hat{\pi}^i$ using (9), then calculating the probabilities as functions of the $\hat{\pi}^i$'s using probabilistic arguments (for example, see Knorr, 1979). Denote the estimates determined by this latter approach by $\tilde{p}_{Z^\bullet}(j)$. Finally, form the squared error quantity

$$E^2(P(A)) = \sum_{j=0}^{k} (\hat{p}_{Z^\bullet}(j) - \tilde{p}_{Z^\bullet}(j))^2 , \tag{12}$$

and find the "least squares" value $0 < \hat{P}(A) < 1$ minimizing (12). This minimization can easily be carried out numerically. Once $\hat{P}(A)$ is obtained, the $\hat{\pi}^i$ are given by (9) with $\hat{P}(A)$ in place of $P(A)$. The total number s of signals present can then be estimated by $\hat{s} = t^\bullet/\hat{P}(A)$.

A variation of this, which might be described as an overconstrained method of moments with least squares fit, is as follows. As in the method of moments, set

$$t^i = \hat{s}\hat{\pi}^i \; ; \; i = 1,2,\ldots,k$$

This gives $2k$ equations in the $k + 1$ unknowns $\hat{s}$, $\hat{\pi}^1,\ldots,\hat{\pi}^k$, since the $p_{z^\bullet}(j)$'s are functions of the π^i's (and we take $\hat{p}_{z^\bullet}(j)$ to be such functions of the $\hat{\pi}^i$'s). We estimate s and π with a weighted least squares, weighting each expression by the inverse of its variance. Now $V(T^i) = s\pi^i(1-\pi^i)$ and $V(I^j) = sp_{z^\bullet}(j)(1-p_{z^\bullet}(j))$, so we wish to find s and π minimizing

$$E^2(s,\pi^1,\ldots,\pi^k) = \sum_{i=1}^{k} \frac{(t^i - s\pi^i)^2}{s\pi^i(1-\pi^i)} + \sum_{j=1}^{k} \frac{\left\{I^j - sp_{z^\bullet}(j)\right\}^2}{sp_{z^\bullet}(j)(1-p_{z^\bullet}(j))} \; .$$

This minimization can be carried out numerically, using the Generalized Reduced Gradient non-linear maximization algorithm, for example.

(c) Maximum Likelihood Estimators. Recall the 2^k-dimensional vector N, whose "z<u>th</u>" component N^z denotes the number of times $z \varepsilon S$ occurs, is distributed multinomial,

$$N \sim M_{2^k}(s,p) \; ,$$

where in turn z is a k-dimensional vector whose components indicate detection/non-detection of a given signal with the corresponding k receivers. Similarly, p is a 2^k-dimensional vector whose "z<u>th</u>" component, p^z, denotes the probability of observing such a detect/non-detect pattern among the receivers, $p^z = P[Z = z]; \; z \varepsilon S$. Note this formulation imposes no condition

on independence of the receivers, although actual computation of the p^z's in terms of individual receiver detection probabilities (the π^i's) would involve such considerations.

Following the development by Sanathanan (1972), the likelihood function L, evaluated at the observation $(n^1, n^2, \ldots, n^{2^k-1})$, a (2^k-1) dimensional vector (where we omit n^0, the number of signals that went undetected, since $z = 0$ is unobservable), is

$$\text{(13)} \qquad L(s,p) = \binom{s}{n^1, n^2, \ldots n^{2^k-1}} (p^1)^{n^1} \cdots \left(p^{2^k-1}\right)^{n^{2^k-1}} \left(p^0\right)^{s-t^\bullet} ,$$

where $s-t^\bullet = n^0$. It is convenient to factor (13) into a marginal likelihood for s times a conditional likelihood for n given s:

$$\text{(14)} \qquad L(s,p) = \binom{s}{t^\bullet} (p^0)^{s-t^\bullet} (1 - p^0)^{t^\bullet}$$

$$\times \binom{t^\bullet}{n^1 \ldots n^{2^k-1}} (c^1)^{n^1} \cdots \left(c^{2^k-1}\right)^{n^{2^k-1}} ,$$

$$= L_1(s,p^0) \times L_2(p) ,$$

where $c^i = p^i/(1-p^0)$. Sanathanan considers both the maximum likelihood estimators $\hat{s}$ and $\hat{p}$ maximizing (13) and "conditional maximum likelihood estimators" $\tilde{s}$ and $\tilde{p}$, where

$\tilde{p}$ maximizes $L_2(p)$ and $\tilde{s}$ maximizes $L_1(s,\tilde{p}^0)$. In this situation, it is known that, once $\tilde{p}$ has been found, the solution for $\tilde{s}$ is given by $\lceil(t^\bullet/(1-\tilde{p}^0)$, where $\lceil(x)$ denotes the greatest integer not exceeding x .

In our model, with p^z given in equation (2), the conditional m.ℓ.e. approach will involve maximization of the second factor in (14) over a (k+m) - dimensional space, where m is the number of unknown parameters in the distribution of signal strength, V . We shall concentrate on two special cases: one in which $\sigma_V^2 = 0$, so V can be considered to be degenerate at 1 ; and one in which V is distributed Beta with parameters $\alpha > 0$ and $\beta > 0$, as in our earlier example. In either case the maximization can be carried out with a numerical minimization routine such as GRG.

Case 1: V degenerate at 1 . The conditional likelihood function $L_2(p)$ is a function of k parameters, $\ell_1,\ldots,\ell_k$, since in this case the receivers are independent and, by (2), $p^z = \prod_{i=1}^{k} \ell_i^{z^i}(1-\ell_i)^{1-z^i}$ for $z \neq 0$. (For $z = 0$, $p^0 = 1 - \sum_{z\neq 0} p^z = 1 - \prod_{i=1}^{k}(1-\ell_i))$. Once the n^z's $(z \neq 0)$ are observed, $t^\bullet = \sum_{z\neq 0} n^z$ is known, and L_2 in (14) is a function of $\ell_1,\ldots,\ell_k$. Note in this case $\pi^i = \ell_i$, so the individual receiver detection probabilities are estimated by the $\tilde{\ell}_i$, and $\tilde{s} = \lceil(\tau^\bullet/\prod_{i=1}^{k}(1-\tilde{\ell}_i))$. The numerical maximization of L_2 in this case is over the k-dimensional space of ℓ_i's .

Case 2: $V \sim \mathrm{Beta}(\alpha,\beta)$. In this case, equation (2') gives

$$(15) \qquad p^z = \frac{1}{B(\alpha,\beta)} \int_0^1 \prod_{i=1}^{n} \left[(\ell_i v)^{z^i} (1-\ell_i v)^{1-z^i} \right] v^{\alpha-1} (1-v)^{\beta-1} \, dv \; .$$

Now for each trial set of ℓ_i's and α , β in the numerical maximization of L_2 in (14), the p^z's can be obtained from (15) by numerical integration. Such numerical integration thus becomes part of the process of evaluating the objective function at each stage of the numerical maximization; this maximization would be over the (k+2) - dimensional space of the ℓ_i's and α, β . As a special case, if $\beta = 1$ and the receivers are identical (the ℓ_i's are the same), the p^z's in equation (15) can be obtained explicitly as an incomplete beta function:

$$p^z = \int_0^1 (\ell v)^{z^\bullet} (1-\ell v)^{k-z^\bullet} \alpha v^{\alpha-1} \, dv$$

$$= \frac{\alpha}{\ell^\alpha} \int_0^\ell u^{(z^\bullet+\alpha)-1} (1-\ell v)^{(k-z^\bullet+1)-1} \, dv$$

$$= \frac{\alpha}{\ell^\alpha} B(\ell;\; z^\bullet+\alpha,\; k-z^\bullet+1) \; ,$$

where $B(x;\, a,b)$ is the incomplete beta function with parameters

a and b. This function is available as a subroutine in many computer installations, so a separate numerical integration routine need not be called in this special case. The numerical maximization of L_2 is over the 2-dimensional space (ℓ,α) in this case.

6. Testing the Model

A chi-square test of a model for the signal detection process can be made in cases where the maximum likelihood approach is used. This can be applied, somewhat more approximately, when conditional maximum likelihood approach is used. This test can be used, for example, to investigate whether the receiver detection probability model ($\tau_i(V) = \ell_i \cdot V$ with $V \sim F$ and "conditional independence" of the receivers) is plausible, in light of the observed data. In particular, with V degenerate at 1, it can be used to test the plausibility of the assumption of receiver independence.

The test is conducted as follows. Once the p^z and s have been estimated, the expected number of outcomes of each type z for $z \varepsilon S$ can be estimated by $\hat{E}(N^z) = \hat{s} \cdot \hat{p}^z$, and the test statistic

$$t = \sum_{z \neq 0} (n^z - \hat{E}(N^z))^2 / \hat{E}(N^z)$$

computed. Under the hypothesized model, t is (approximately) an outcome from a chi-square population with $(2^k-1) - (m+1)$

degrees of freedom, where m is the number of parameters estimated (i.e., the dimension of the space over which the numerical maximization was conducted). As usual for chi-square "goodness-of-fit tests, if some of the $\hat{E}(N^z)$'s are less than three or so, these "categories" should be pooled with others so the three or more expected count rule of thumb is met, with corresponding reduction in the degrees of freedom for the test statistic. The level of significance of the test is the tail area above t, under the appropriate chi-square density.

7. Comparison of the Estimators by Simulation

CAPT Dave Hendrickx is currently conducting a series of simulations to evaluate and compare the behavior of the estimators described above, for a variety of data input streams related to the general model $\tau_i(V) = \ell_i \cdot V$, with $V \sim \text{Beta}(\alpha,1)$. By varying α from 1 to 100, varying degrees of intraclass correlation can be generated among the simulated receivers. As mentioned above, with $\alpha \approx 100$, the receivers are essentially independent, since V is then nearly degenerate at 1. On the other hand, with $\alpha = 1$ the signal strengths vary at random over $(0,1)$, and there is appreciable intraclass correlation among the receivers. Details of the simulation and analyses of the results will be published in Hendrickx's masters thesis, Hendrickx (1981).

Preliminary analyses of results from a limited number of simulation cases suggest the following conjectures concerning the three estimation procedures:

1) it is not the case that one of the procedures dominates the others over all situations;

2) generally speaking, the maximum likelihood estimator has smaller variance than does the method of moments or least squares;

3) all procedures appear to underestimate s and the π_i's when the receivers are made dependent by taking α as small as 1; and

4) the estimators perform generally surprisingly well, even when the total number of signals s is as small as 25 or so, without regard to the number k of receivers or the mix of π_i's, when the receivers are (at least nearly) independent.

As remarked above, these are preliminary observations; more details will be documented by Hendrickx.

REFERENCES

[1] Cormack, R. M. (1972), The Logic of Capture-Recapture Estimates, Biometrics, 28, 337-343.

[2] Sanathanan, Lalitha (1972), Estimating the Size of a multinomial Population, The Annals of Mathematical Statistics, 43, 142-152.

[3] Knorr, Jeffrey B. (1979), A Model for Estimating Acquisition Receiving System Intercept Probability, Naval Postgraduate School Technical Report NPS62-79-005, 29p.

[4] Scheffe, Henry (1959), The Analysis of Variance, New York, John Wiley & Sons, Inc.

[5] Lasdon, L., A. Waren, A. Jain and M. Ratner (1975), GRG System. Documentation, Cleveland State University Technical Memorandum No. CIS-75-01.

[6] Hendrickx, David (to be published in 1981), An evaluation of estimators for receiver detection probabilities, Masters Thesis in Operations Research, Naval Postgraduate School.

Appendix

LIST OF SYMBOLS
(in order of occurrence)

k: number of receivers

s: number of signals

π_{ij}: probability receiver i detects signal j

V_j: strength of signal j ; $V_j \sim F(v)$, $E(V) = \mu_V$, $V(V) = \sigma_V^2$.

$\tau_i(v_j) = \pi_{ij} = \ell_i \cdot v_j$

$Z = (Z^1,\ldots,Z^k)$: vector of Bernoulli detect/non-detect indicators, for a given signal, for all receivers.

$\pi^i = P[Z^i = 1] = E(\tau_i(V)) = \ell_i \cdot \mu_V$.

S: sample space of Z .

$p^z = P[Z=z]$; $z \varepsilon S$.

n^z: number of times $z \varepsilon S$ occurs

$n = (n^0, n^1, \ldots, n^{2^k-1})$: vector of n^z's in some specified order; $n^z|_{z=0} = n^0$

$p = (p^0, p^1, \ldots, p^{2^k-1})$: vector of p^z's in same order as for n .

N: random vector with outcome n ; $N \sim M_{2^k}(s,p)$

$R(j)$: the jth residual; $R(j)$ = (number of signals detected by j receivers) − (expected number of signals detected by j receivers).

T^i: number of signals detected by receiver i (outcome is t^i)

T^*: $\sum_{i<j}$ (number of signals detected simultaneously by receivers i and j)

A: the event $[Z \neq 0]$

$\hat{s}$, $\hat{\pi}^i$: estimates of s and π^i , respectively

$t^{\bullet}$: number of signals detected

$Z^{\bullet} = \sum_i Z^i$: number of detections of a given signal

$p_{Z^{\cdot}}(j)$: mass function of $Z^{\cdot}$

I^j: number of data vectors z for which $z^{\cdot} = j$

$L(s,p)$: likelihood of N at n

$c^i = p^i/(1-p^0)$

DISTRIBUTION LIST

	NO. OF COPIES
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Library, Code 0142 Naval Postgraduate School Monterey, CA 93940	2
Dean of Research Code 012 Naval Postgraduate School Monterey, CA 93940	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93940	1
Prof. D. R. Barr Code 55Bn Naval Postgraduate School Monterey, CA 93940	69